Session 5pAA

Architectural Acoustics, Speech Communication and Psychological and Physiological Acoustics: Psychological Aspects of Speech in Rooms II

John S. Bradley, Cochair
National Research Council, Inst. for Research in Construction, Ottawa, Ontario K1A 0R6, Canada

Hiroshi Sato, Cochair
National Inst. for Advanced Industrial Science Technology, Inst. for Human Science and Biomedical Engineering, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566 Japan

Contributed Papers

1:00

5pAA1. Calibration of consonant perception in room reverberation. Kanako Ueno (Inst. of Industrial Sci., Univ. of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505, Japan), Norbert Kopco, and Barbara Shinn-Cunningham (Boston Univ., 677 Beacon St., Boston, MA 02215)

Many studies of sound perception often assumed that our auditory sensory processes are relatively static, rather than plastic. However, in everyday environments, we naturally and fluidly compensate for interfering effects of background noise and room reverberation. In order to investigate how listeners calibrate auditory perception to such acoustic interference, a listening experiment was performed to measure the effect of sudden changes of reverberation on the identification of consonants. Test sounds were generated by convolving two types of binaural room impulse responses (BRIRs) measured in large real rooms with speech tokens. As a control condition, pseudo-anechoic BRIRs with negligible reverberation energy were used. Listeners were asked to identify the consonant in a vowel-consonant target. The target was preceded by a carrier phrase consisting of vowel-consonant pairs from the same talker. In some cases, the target and carrier phrase were processed by the same BRIRs, while in others the BRIR's processing target and carrier differed. Consistent effect of calibration was observed in one of the simulated rooms, but not in the other, suggesting that the ability to compensate for the effects of reverberation depends on the specific pattern of reverberation produced in a given room. [Work supported by AFOSR and NSF.]

1:15

5pAA2. Sentence context influences vowel perception in reverberant conditions. Janine Wotton (Dept. of Psych., Gustavus Adolphus College, 800 W. College Ave., St. Peter, MN 56082, jwotton2@gac.edu), Kristin Welsh, Crystal Smith, Rachel Elvebak, Samantha Haseltine (Gustavus Adolphus College, St. Peter, MN 56082), and Barbara Shinn-Cunningham (Boston Univ., Boston, MA 02115)

Sentences recorded with a Mid-western accent were convolved with head-related impulse responses that included different room reverberation conditions. The stimuli were presented binaurally through headphones in an echo-attenuated chamber and subjects ($n=23$) typed the sentences they heard. The target word was one of a vowel pair (cattle/kettle, jam/gem, gas/guess, past/pest) embedded as the second word in one of three sentence types. The neutral sentence provided little context for the word. Target words in sentences that provided strong contextual cues could be congruent or incongruent with the expectations of the subject, for example, ''The cattle/kettle grazed in the meadow.'' Subjects made significantly more errors in the incongruent sentences compared to the neutral (Wilcoxon$=3.572$ $p<0.05$) or congruent sentences (Wilcoxon$=3.56$ $p<0.05$). When the target word was in a congruent sentence subjects performed equally well in reverberant or pseudo-anechoic conditions (Wilcoxon$=1.298$) but they made more errors in the reverberant condition for both neutral (Wilcoxon$=3.359$, $p<0.05$) and incongruent sentences (Wilcoxon$=2.241$, $p<0.05$). Results suggest that reverberation may cause listeners to rely more heavily on linguistic context to determine word meaning. [Work supported by NOHR, AFOSR.]

1:30

5pAA3. Perceptual compensation for reverberation: Effects of noise-context bandwidth. Simon J. Makin, Anthony J. Watkins, and Andrew P. Raimond (School of Psych., The Univ. of Reading, Earley Gate, Reading RG6 6AL, UK, s.j.makin@reading.ac.uk)

Perceptual compensation for reverberation is observed when the reverberation is applied to a test word (from a ''sir''-to-''stir'' continuum) and its context (e.g., ''next you'll get_to click on'') are varied independently. Increasing reverberation in test words decreases listeners' ''stir'' responses, as reverberation ''fills the gap'' that cues the [t]. Compensation occurs when the context's reverberation is commensurately increased, and ''stir'' responses increase back to the level found with minimal test-word reverberation. Compensation is strongest with speech contexts but also occurs with some noise-like contexts, including ''signal-correlated noise'' that has the wideband temporal envelope of the original speech. Also effective is a wideband noise that is given the temporal envelope seen at the output of a single auditory filter in response to speech. A narrow-band version of this ''auditory-filtered'' noise is not effective, but when contexts are made by summing of three or five of these bands, their effectiveness increases correspondingly. Compensation appears to be informed by the ''tails'' that reverberation adds at offsets, so it merely requires contexts with suitable temporal-envelope fluctuations. However, effects seem confined to the context's frequency region, so the crucial offsets need to be in a wide range of frequency bands. [Work supported by EPSRC.]

1:45

5pAA4. Aural localization of speech stimuli. Evelyn Way (Talaske, 105 N. Oak Park Ave, Oak Park, IL 60301, evelyn@talaske.com)

Localization error was tested for a variety of signals to answer the question: do humans aurally localize different speech stimuli with a different localization blur? A series of tests was conducted comparing the effect of the sentence length, gender of the talker, and frequency content of the signals on localization. Results were applied to ongoing research into constructing an aurally accurate telepresence system.

2:00

5pAA5. Quantifying the effects of room acoustics on speech intelligibility in very low signal-to-noise ratio environments. Jarrod E. Whittington (Grad. Program in Architectural Acoust., Rensselaer Polytechnic Inst., Troy, NY 12180, whittj@rpi.edu) and John S. Bradley (Inst. for Res. in Construction, Ottawa, K1A OR6 Canada)

The intelligibility of speech transmitted from closed offices to adjacent spaces is strongly affected by the signal-to-noise ratio at the receiver position and the acoustical characteristics of the spaces involved. Previous studies have suggested that the effect of room acoustics on speech intelligibility in closed offices and rooms is negligible and can be ignored (as with intelligibility quantifiers such as the articulation index). The purpose of this study is to show that in conditions of very low signal-to-noise (i.e., when high speech privacy is a primary concern), the influence of room acoustics rises dramatically. To this end, multiple subjects were given tests of speech intelligibility in simulated sound fields. Speech samples were presented to subjects with seven levels of signal-to-noise and four different reverberation times. The results from these tests show that as reverberation time increases, speech intelligibility decreases much more sharply for very low (-8 dB) signal-to-noise situations than in higher ($+10$ dB) signal-to-noise situations. This suggests an important relationship between room acoustics and speech privacy/security.

2:15

5pAA6. The effect of a preprocessing approach improving speech intelligibility in reverberation considering a public-address system and room acoustics. Nao Hodoshima, Takayuki Arai (Dept. of Elec. and Electron. Eng., Sophia Univ., Tokyo 102-8554, Japan), and Peter Svensson (Norwegian Univ. of Sci. and Technol., Trondheim NO-7491, Norway)

This study evaluates a preprocessing approach for reducing reverberation effects when the input signal is not ideal, dry speech, but rather a realistic speech signal picked up by a close microphone in a room. And this study shows how the situation affects the input and the further approach compared to a dry signal as the input. Steady-state suppression, as described by Arai *et al.* [Acoust. Sci. Technol. **23**, 229–232 (2002)], was used as a preprocessing approach that processes a speech signal before it is radiated from loudspeakers. A lecture was simulated in two different halls (reverberation times of 1.2 and 1.8 s) in which public address systems were installed. The simulation software CATT-Acoustic was used and impulse responses were calculated for the input to the preprocessing approach and for a listener position. Stimuli for a syllable identification test were prepared by convolving speech signals with the calculated impulse responses. Speech signals were given with and without steady-state suppression. The inclusion of natural and electroacoustical impulse responses makes the study of steady-state suppression more realistic and tests its robustness. [Work supported by JSPS (176911).]

2:30

5pAA7. Improving speech intelligibility in reverberant rooms. Douglas F. Winker and Elmer L. Hixson (Dept. of Elec. and Comp. Eng., Univ. of Texas, Austin, TX 78712)

Speech intelligibility in rooms with long reverberation times has long been a problem for acousticians. This problem is exacerbated when the option for acoustic treatments is not an option for a variety of reasons. In this study, a constant-beamwidth, wide-bandwidth (CBWB) loudspeaker array was used to improve speech intelligibility in two multipurpose rooms where treatment was not an option. A CBWB loudspeaker array was designed with independent beam pattern control on both the horizontal and vertical axes. A wide horizontal beam pattern and narrow vertical pattern were designed to achieve constant coverage from 500 Hz to 4 kHz. The array incorporates four separate arrays, two with nested elements, and linear-phase FIR filters to maintain constant beamwidth over the three octaves of interest. Room simulations were conducted with CATT-Acoustic and the array was implemented and compared to a source with a more conventional coverage pattern. The STI method was used to measure speech intelligibility at several positions throughout the rooms. The study showed an improvement in speech intelligibility and a reduction in measured reverberation times.

2:45

5pAA8. Experimental investigations of the influence of room acoustics on the teacher's voice. Malte Kob, Anja Kamprolf, Christiane Neuschaefer-Rube, Oliver Goldschmidt (Chair of Phoniatrics and Pedaudiology, RWTH Aachen Univ., Pauwelsstr. 30, 52074 Aachen, Germany, mkob@ukaachen.de), and Gottfried Behler (RWTH Aachen Univ., 52066 Aachen, Germany)

Teachers belong to the group of professional voice users who often suffer from voice disorders. One reason for a significantly increased prevalence of voice problems can be poor room acoustical conditions in the classrooms. In this study, four rather reverberant and loud classrooms in a primary school in Aachen were analyzed using measurements of the reverberation time, T_{20}, and the speech transmission index, STI. About half of the school's teachers were investigated with respect to their voice status by using phoniatric, logopedic, and objective voice analysis methods. The prevalence of voice problems in this group was found to exceed previous studies where subjective voice quality was rated. In a second part of this joint project the change of voice quality during the teachers' working day was analyzed. Two of the four rooms were acoustically optimized. Members of two groups of teachers with and without voice problems were recorded before and after teaching in either one of the acoustically poor rooms or one of the newly renovated rooms. The preliminary results indicate changes of the voice quality in most subjects with respect to one or more voice parameters. Further studies shall prove the significance of the room influence.

3:00–3:15 Break

3:15

5pAA9. Human voice phoneme directivity pattern measurements. Brian F. G. Katz, Fabien Prezat, and Christophe d'Alessandro (Percept. Située, LIMSI-CNRS, BP 133, F91403 Orsay, France)

The application of directivity patterns to radiating sources into computer simulations and auralizations is common for loudspeaker models. Few applications include the directivity patterns of natural sources, partly due to the lack of sufficient data. This work presents the results of a detailed measurement study on human voice directivity in three-dimensions. Unlike previous studies that have used average directivity data over read phrases, this work presents results that are measured for a number of sustained individual phonemes. Details of the measurement protocol and posttreatment processing are presented. Comparisons are made relative to phoneme, f0, spectral characteristics, and associated mouth geometry for several talker subjects. Studies have also been made on the directivity of the singing voice. Specifically, the variations in directivity relative to level (piano, fortissimo, etc.) and projection as controlled by the singer have been investigated. Results of this work are applicable to speech production research, talker simulator design, room acoustic sound field prediction, and virtual reality systems with talking avatars.

3:30

5pAA10. Case study for voice amplification in a highly absorptive conference room using a negative absorption tuning by the Yamaha Active Field Control system. Takayuki Watanabe (YAMAHA Commercial Audio Systems, Inc., 6600 Orangethorpe Ave., Buena Park, CA 90620)

The Yamaha Active Field Control (AFC) enhances and varies acoustical conditions from the room acoustics using the acoustical feedback in a system, which is properly tuned to the installed room acoustics. The system is primarily used to improve auditory impressions of reverberance, loudness by enhancement of reverberation in time, and energy in a music venue. Here, a negative absorption tuning is demonstrated as beneficial to